U.S. PATENT APPLICATION

for

HIGH DENSITY CONTACT TO RELAXED GEOMETRY LAYERS

Inventors:

Roy E. Scheuerlein Christopher J. Petti

Prepared by:
Pamela J. Squyres

Matrix Semiconductor, Inc.
3230 Scott Boulevard

Santa Clara CA 95054

HIGH DENSITY CONTACT TO RELAXED GEOMETRY LAYERS

RELATED APPLICATIONS

[0001] This application is related to Cleeves et al., US Application No. ______, "Optimization of Critical Dimensions and Pitch of Patterned Features in and Above a Substrate," (attorney docket no. MA-110) filed on even date herewith, which application is hereby incorporated by reference.

BACKGROUND OF THE INVENTION

[0002] The invention relates to a structure for providing electrical interconnection between levels formed at different heights in an integrated circuit, specifically to elements in a dense multilevel array.

[0003] It is known to form vias to electrically connect routing levels formed at different heights in an integrated circuit. In a conventional arrangement, shown in Fig. 1, bottom routing level RL_1 is formed first. Generally a dielectric material is deposited on RL_1 . The dielectric material is selectively excavated to form a hole, which is then filled with a conductive material to form via V_1 . In the same manner routing level RL_2 is formed above via V_1 , via V_2 formed above routing level RL_2 , and so on, alternating routing levels and vias providing electrical interconnection as required.

[0004] With the advent of extremely dense structures requiring extensive interconnection, including, for example, interconnection between layers that are not vertically adjacent, while the highest possible density must be maintained, this conventional arrangement uses space inefficiently and requires extra processing steps.

[0005] There is a need, therefore, to form vias between different levels in a dense array while using as little die area as possible.

SUMMARY OF THE INVENTION

[0006] The present invention is defined by the following claims, and nothing in this section should be taken as a limitation on those claims. In general, the invention is directed to a novel structure of vias and staggered routing levels adapted to provide electrical connection to a dense multilevel array.

[0007] A first aspect of the invention provides for a structure for providing multilevel electrical connectivity within an integrated circuit, the structure comprising a first plurality of vias; a second plurality of vias, wherein the first and second pluralities of vias are vertically overlapping; a first routing level at a first height, said first level connected to the first plurality of vias; and a second routing level at a second height, said second level connected to the second plurality of vias, wherein the first height is different from the second height, wherein both routing levels are formed above the substrate, and wherein a) the first routing level and the second routing level are above the first and second vias or b) the first routing level and the second routing level are below the first and second vias.

[0008] Another aspect of the invention provides for a structure for providing multilevel electrical connectivity within an integrated circuit, the structure comprising a first plurality of vias; a second plurality of vias, wherein the first and second pluralities of vias are vertically overlapping; a first routing level at a first height, said first level connected to the first plurality of vias; a second routing level at a second height, said second level connected to the second plurality of vias, wherein the first height is different from the second height; and a third routing level at a third height, the third level connected to the first plurality of vias and to the second plurality of vias, wherein all three routing levels are formed above the substrate.

[0009] A related aspect of the invention provides for a structure for providing multilevel electrical connectivity within an integrated circuit, the structure comprising a first plurality of vias; a second plurality of vias; a first routing level at a first height, said first level connected to the first plurality of vias; a second routing level at a second

height, said second level connected to the second plurality of vias, wherein the first height is different from the second height; and a third routing level at a third height, the third level connected to the first plurality of vias and to the second plurality of vias, wherein all three routing levels are formed above the substrate, and wherein either the third routing level is above both the first and the second routing levels, or the third routing level is below both the first and the second routing levels.

[0010] Yet another aspect of the invention provides for a method for forming a via and routing structure for electrically connecting a multilevel array in an integrated circuit, the method comprising forming a first routing level; forming a second routing level above the first routing level; forming a first plurality of vias connected at bottom ends to the first routing level; forming a second plurality of vias connected at bottom ends to the second routing level, wherein the first and second pluralities of vias are vertically overlapping.

[0011] Each of the aspects and embodiments of the invention can be used alone or in combination with one another.

[0012] The preferred aspects and embodiments will now be described with reference to the attached drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] Fig. 1 is a cross-sectional view of a conventional routing-level-and-via arrangement used in prior art.

[0014] Fig. 2 is a cross-section illustrating relative pitch of vias and routing level when the vias all attach to a single routing level.

[0015] Fig. 3. is a cross-section of vias and staggered routing levels according to the present invention, staggered routing levels formed below the vias.

- [0016] Fig. 4. is a cross-section of vias and staggered routing levels according to the present invention, staggered routing levels formed above the vias.
- [0017] Fig. 5 is a plan view of a memory level with memory lines connected to vias according to the present invention.
- [0018] Fig. 6 is a cross-section of vias shown Fig. 5, their attachment to staggered routing levels illustrated.
- [0019] Fig. 7 is a cross-section showing the use of passing lines at the height of the first routing level.
- [0020] Figs. 8a through 8f are cross-sections illustrating fabrication of vias and staggered routing levels according to the present invention.
- [0021] Figs. 9a through 9c are cross-sections illustrating an alternate method of fabrication of vias and staggered routing levels according to the present invention.
- [0022] Figs. 10a through 10d are cross-sections illustrating fabrication of vias and staggered routing levels according to the present invention with the staggered routing levels formed above the vias.
- [0023] Figs. 11a and 11b are cross-sections illustrating an alternate method of fabrication of vias and staggered routing levels according to the present invention with the staggered routing levels formed above the vias.

DETAILED DESCRIPTION OF THE INVENTION

[0024] Monolithic three dimensional memory arrays are described in Johnson et al., US Patent No. 6,034,882, "Vertically stacked field programmable nonvolatile memory and method of fabrication"; Johnson, US Patent No. 6,525,953, "Vertically stacked field programmable nonvolatile memory and method of fabrication"; Knall et al., US Patent

No. 6,420,215, "Three Dimensional Memory Array and Method of Fabrication"; Lee et al., US Patent Application No. 09/927648, "Dense Arrays and Charge Storage Devices, and Methods for Making Same," filed August 13, 2001; Herner, US Application No. 10/095962, "Silicide-Silicon Oxide-Semiconductor Antifuse Device and Method of Making," filed March 13, 2002; Vyvoda et al., US Patent Application No. 10/185507, "Electrically Isolated Pillars in Active Devices," filed June 27, 2002; Herner et al., US Patent Application No. 10/326470, "An Improved Method for Making High Density Nonvolatile Memory," filed Dec. 19, 2002; Walker et al., US Application No. 10/335089, "Method for Fabricating Programmable Memory Array Structures Incorporating Series-Connected Transistor Strings," filed December 31, 2002; Scheuerlein et al., US Application No. 10/335078, "Programmable Memory Array Structure Incorporating Series-Connected Transistor Strings and Methods for Fabrication and Operation of Same," filed December 31, 2002; and Vyvoda, US Patent Application No. 10/440882, "Rail Schottky Device and Method of Making", filed May 19, 2003, all assigned to the assignee of the present invention and hereby incorporated by reference.

[0025] A monolithic three dimensional memory array is one in which multiple memory levels are formed above a single substrate, such as a wafer, with no intervening substrates. In contrast, stacked memories have been constructed by forming memory levels on separate substrates and adhering the memory levels atop each other, as in Leedy, US Patent No. 5,915,167, "Three dimensional structure memory." The substrates may be thinned or removed from the memory levels before bonding, but as the memory levels are initially formed over separate substrates, such memories are not true monolithic three dimensional memory arrays.

[0026] In some of these memory arrays, for example the array of Scheuerlein et al., the memory cells are charge storage memory cells, like SONOS or floating gate devices. In others, such as Herner et al., the memory cells are write-once memory cells, for example a paired antifuse and diode, the cell in one programmed state before the antifuse is ruptured, and in another programmed state after.

[0027] In embodiments of these monolithic three dimensional memory arrays, dense memory lines must be electrically connected to underlying or overlying interconnect layers. These interconnect levels may sometimes be shared, and need not be as dense as the memory levels. In a repeating pattern, such as a line-and-space pattern, *pitch* is the distance from one point in the pattern to the next occurrence of the same point: the distance from the center of one line to the center of the adjacent line, for example. It is advantageous for the memory to be as dense as possible, with the tightest possible pitch. Since interconnect layers need not be as dense, it is preferable to form them more widely spaced, formed at a lower pitch, using cheaper lithography tools and photomasks and lowering the overall cost of the integrated circuit.

[0028] Difficulties are encountered, however, when trying to electrically connect a level formed at a small pitch to a level formed at a larger pitch.

[0029] Turning to Fig. 2, it is usual for the region where a via makes contact to a routing level R₁ to be a pad which is larger than the minimum feature size F for routing level R₁. (The minimum feature size is the smallest patterned feature or patterned gap that can be formed using given photolithography equipment.) The purpose of the pad is to avoid misalignment and to maximize contact area (thus minimizing resistance.) The pad 10 is typically larger than the feature size on all sides by a tolerance T, which is normally about ½ F. The gap between pads 10 cannot be less than F, the minimum feature size.

[0030] It will be seen that the pitch is P_R for the routing level R_1 . Since there is a one-to-one correspondence between vias 20 and routing level R_1 pads 10, it will be seen that the via pitch P_V is identical to routing level pitch P_R . In this case $P_V=P_R=(T+F+T+F)=3F$, assuming T is $\frac{1}{2}F$.

[0031] Fig. 3 shows an alternative arrangement according to the present invention. Instead of a single routing level R_1 , there are two routing levels, R_1 and R_2 , formed at different heights. In this drawing, vias 20 connect to routing level R_1 , while vias 30 connect to routing level R_2 . Vias 20 alternate with vias 30 and are vertically overlapping.

Two or more vias are considered to be "vertically overlapping" when some portion of each of the vias excluding the routing levels exist at substantially the same height above the substrate. Since pads 10 connecting to adjacent vias 20 and 30 are no longer at the same level, it is not necessary for a space of feature size F to exist between them. It will be seen that with routing levels R_1 and R_2 at different heights, via pitch P_V can be significantly smaller than in Fig. 2. A tolerance should be maintained between vias 20 and adjacent pads 10. This tolerance is not exactly the same as T, since it has a slightly different physical requirement (T´ must be adequate to maintain a minimum dielectric thickness between the contact and the neighboring pad), but is similar, and will be referred to as T´; this discussion assumes that T and T´ are about the same. Thus assuming vias 20 and vias 30 have a feature size of about F, the routing level minimum feature size, and assuming T (and T´) is about $\frac{1}{2}$ F, via pitch P_V is about F+T+T´, or about 2F. This is a 3:2 improvement in via pitch P_V over the arrangement shown in Fig. 2.

[0032] This example showed vias connecting to routing levels R_1 and R_2 , where R_1 and R_2 are formed below vias 20 and 30, and where the pitch of routing levels R_1 and R_2 is larger than the pitch of the structures above them. This is in contrast to the usual trend in integrated circuits, in which pitch is progressively more relaxed the higher a level is formed above the substrate, normally a semiconductor wafer. Such an arrangement is advantageous when, for example, vias 20 and 30 are connecting to memory lines in a memory array, the memory array formed over a device level formed in the substrate, the substrate device level including, for example, CMOS circuits formed in the substrate, when the substrate device level is formed at a larger pitch than the memory array. Such an arrangement is fully described in the related application of Cleeves et al., filed on even date herewith. Routing levels R_1 and R_2 of this example are advantageously formed below the memory array and above the substrate, where connectivity from routing levels R_1 and R_2 to substrate circuitry is easily achieved.

[0033] The staggered routing levels of the present invention would also provide advantage in a more conventional integrated circuit in which pitch increases in levels

formed successively higher above the substrate. In this case the vias would connect to staggered routing levels R_1 and R_2 formed above the vertically overlapping vias, as in Fig. 4.

[0034] A detailed example will be provided of vias connecting memory lines in a monolithic three dimensional memory array to staggered routing levels below the array, the vias and routing levels formed according to the present invention. Further information will be provided as to advantageous methods and materials which are used to fabricate such an array. As will be apparent to those skilled in the art, this detailed example is just one of many possible embodiments of the present invention. The arrangement and placement of the vias and staggered routing levels, the circuits they serve to interconnect, and many details of their formation and use can be widely varied while the result still falls within the scope of the invention.

STRUCTURE

[0035] In general, the monolithic three dimensional memory arrays described in the incorporated applications and patents describe dense memory arrays comprising memory cells of various types, some rewritable, others write-once. The vias and staggered routing levels according to the present invention could be used with any of these monolithic three dimensional memory arrays. Two will be selected for particular discussion.

[0036] In the rewritable memories of Scheuerlein et al., memory cells are charge storage memory cells, for example SONOS or floating gate devices. In such array, lines in one dimension are channel bodies arranged in a NAND series string; substantially perpendicular to these are control gate electrodes.

[0037] Alternatively, as in the memory of Herner et al., the array can be a passive element memory array. In such an array, the lines comprise conductors. At the intersections of the conductors are passive element memory cells. As used herein, a passive element memory array includes a plurality of two-terminal memory cells, each connected between an associated X-line and an associated Y-line. Such a memory array

may be planar or may be a three-dimensional array having more than one plane of memory cells. Each such memory cell has a non-linear conductivity in which the current in a reverse direction (i.e., from cathode to anode) is lower than the current in a forward direction. Application of a voltage from anode to cathode greater than a programming level changes the conductivity of the memory cell. The conductivity may increase when the memory cell incorporates a fuse technology, or may decrease when the memory cell incorporates an antifuse technology. A passive element memory array is not necessarily a one-time programmable (i.e., write once) memory array.

[0038] Thus lines in a monolithic three dimensional memory array may be, for example, the NAND strings of Scheuerlein et al. or the conductors (or bitlines) of Herner et al. In either case, these lines require connection to vias which in turn connect to support circuitry. Connections of the lines to vias are advantageously arranged as in Fig. 5, which shows a level of memory in plan view. In this level, memory lines have a first pitch P_M. Adjacent memory lines, such as lines 202 and 204, are interleaved, connecting to vias at opposite ends; i.e. line 202 connects to via 210, while line 204 connects to via 218. Memory lines on opposite sides of the same via, for example lines 202 and 203 on opposite sides of via 210, share the via. Vias are arranged in rows, here rows 220 (including vias 208, 210, 212, and 214) and 230, each row having a via pitch P_V. The combination of interleaving lines and sharing vias allows every memory line to connect to a via, even though the memory pitch P_M is significantly smaller than the via pitch P_V.

[0039] Fig. 6 shows vertically overlapping vias 208, 210, 212, and 214 from Fig. 5 in cross-section. These vias connect at their bottom ends to staggered routing levels R_1 and R_2 according to the present invention, to pads 240 for routing level R_1 and to pads 250 for routing level R_2 . At their top ends they connect to routing level R_3 at pads 260. As shown in Fig. 5, routing level R_3 is made up of memory lines in a memory array. Routing level R_3 is formed above the vias, vertically opposite the first and second routing levels.

[0040] Note that the minimum feature size in the memory level shown in Fig. 5 is significantly smaller than the minimum feature size in routing levels R_1 and R_2 of Fig. 6.

It is assumed that cheaper lithography tools and cheaper photomasks are used to form levels R_1 and R_2 than to form the dense memory lines of Fig. 5, reducing overall die cost. Thus the memory level pads 260 can be smaller than the routing level R_1 and R_2 memory level pads 240 and 250, as can the gap separating them.

[0041] Turning to Fig. 7, additional connectivity, for example to circuitry formed in the substrate, can be provided by forming passing lines 40 between routing level R_1 pads 10. These lines can be formed at the minimum feature size F and with a gap of minimum feature size F separating them from pads 10. If passing lines are used, via pitch P_V will be about 2.5 times the minimum feature size F.

FABRICATION

[0042] The structures described above having vias connected to staggered routing levels can be formed using any methods known in the art. A detail description of preferred methods to form embodiments of the present invention will be provided.

[0043] BOTTOM ROUTING LEVELS, SINGLE ETCH AND FILL: As noted, the staggered routing levels may be above the vertically overlapping vias or below them. Either style preferably begins with a suitable substrate 300, typically a monocrystalline silicon wafer, shown in cross-section in Fig. 8a. In preferred embodiments, support circuitry, for example CMOS circuit 302, is fabricated in the substrate. The substrate circuits are isolated with dielectric 304, for example silicon dioxide, then a conductive layer or layers are formed on dielectric 304. In the example shown, layer 306 is TiN and layer 308 is tungsten. Tungsten layer 308 will form the bulk of the first routing level, and TiN acts as an adhesion layer. Many other materials can be used, and layers 306 and 308 can be replaced with one, two, or more layers.

[0044] As shown in Fig. 8b, layers 308 and 306 are patterned and etched to form first routing level pads 240. If passing lines are being formed, they will be patterned and etched at the same time. In Fig. 8c, dielectric material 304 fills the spaces between pads 240 and covers them to a second height and is planarized, for example by etchback or

chemical mechanical polishing. Conductive materials are deposited, patterned and etched in the same manner as pads 240 to form second routing levels pads 250.

[0045] Next memory levels are constructed. As each level is constructed, more dielectric 304 is deposited to isolate structures, and is planarized after each deposition. Turning to Fig. 8d, when a memory level is to be formed that is intended to be connected by vias to first routing level pads 240 and second routing level pads 250, fabrication of that memory level begins by patterning and etching via holes 310 and 312. Via holes 310 reach to first routing level pads 240, while via holes 312 are shallower, reaching only to second routing level pads 250. It will be recalled that a preferred material for dielectric 304 is silicon dioxide, while the preferred material forming the top of pads 240 and 250 was tungsten (layer 308 in Fig. 8a.) Silicon dioxide and tungsten are advantageous choices in that they have very good etch selectivity; i.e. etchants are available that will etch silicon dioxide at a much higher rate than tungsten. It is important to select materials for these layers for good etch selectivity, since the etch of shallower via holes 312 must be relied on to substantially stop while the etch of deeper via holes 310 continues.

[0046] Turning to Fig. 8e, in this exemplary embodiment a conformal layer of TiN 314 is deposited lining via holes 310 and 312 and forming a thin layer of TiN on dielectric 304. Conductive layer 316, for example of tungsten, is deposited on layer 314, filling via holes 310 and 312 (not shown, as they are now filled) and forming a conductive layer at the memory level. As shown in Fig. 8f, layers 314 and 312 are patterned and etched to form memory level pads 260. Memory lines in the memory level can be formed in the same patterning and etch steps that form pads 260. Pads 240, 250, and 260 of Figs. 8b through 8f correspond to pads 240, 250, and 260 of Fig. 6.

[0047] In this example, the first, deeper vias were connected at bottom ends to the first routing level (pads 240), and the second, shallower vias were connected at bottom ends to the second routing level (pads 250.) The first via holes 310 and second via holes 312 were etched in the same etch process, and the first via holes 310 and second via holes 312 were filled at substantially the same time.

[0048] BOTTOM ROUTING LEVELS, SPLIT ETCHES AND FILLS: As those skilled in the art will appreciate, many other fabrication methods and materials can be used to form the via and staggered routing level structure of Fig. 8f while still falling within the scope of the present invention. For example, turning to Fig. 9a, first routing level pads 240 can be formed as described previously. Next dielectric fill 304 is deposited to the height of the second routing level, then etched to form bottom portions of vias holes reaching to first routing level pads 240. The via holes are filled with conductive material 28, preferably doped polycrystalline silicon, herein called polysilicon, forming bottom portions of vias 20. After planarization, conductive material, preferably titanium nitride 306 and tungsten 308, is deposited on dielectric fill 304 and bottom portions of vias 20. The resulting structure is shown in Fig. 9a.

[0049] As shown in Fig. 9b, layers 306 and 308 are patterned and etched to form second routing level pads 250. A thin etch stop layer 22, for example silicon nitride, is formed over and between pads 250.

[0050] Dielectric fill 304 is deposited between and over pads 250 and etch stop layer 22. Memory levels are constructed. As each level is constructed, more dielectric fill 304 is deposited to isolate structures, and is planarized after each deposition. When a memory level is to be formed that is intended to connect by vias to first routing level pads 240 and second routing level pads 250, via holes reaching to second routing level pads 250 and the tops of vias 20 are patterned and etched. Via holes are etched using a two-step etch process. First a silicon dioxide etch forms the via holes and stops on silicon nitride etch stop layer 22. Next silicon nitride etch stop layer 22 is etched, stopping on the silicon 28 of vias 20 and, in the case of any misalignment, silicon dioxide fill 304, and the tungsten of pads 250. Conductive material, preferably a layer of titanium nitride and a layer of tungsten, is deposited, filling top portions of the via holes and forming a layer above the dielectric, which is then patterned and etched to form third routing level pads 260. The gaps between pads 260 are filled with dielectric fill 304, as shown in Fig. 9c.

[0051] In this example, the first, deeper vias were connected at bottom ends to the first routing level (pads 240), and the second, shallower vias were connected at bottom

ends to the second routing level (pads 250.) The top portions of the deeper via holes and the shallower via holes were etched in the same etch process. The top portions of the deeper vias holes and the shallower via holes were filled at substantially the same time. Compared to previous embodiments, this embodiment requires an extra masking step and an extra contact fill process, but has less stringent etch selectivity requirements.

[0052] TOP ROUTING LEVELS, SPLIT ETCHES AND FILLS: Various fabrication methods can be used to form vias for embodiments in which the staggered routing levels attach to top ends of the vias, as in Fig. 4. Examples will be given of a method to form vias attaching to staggered first and second routing levels formed above the vias, and, at opposite ends of the vias, to a third routing level formed below the vias. As described earlier, the third routing level may comprise memory lines in a memory array. Turning to Fig. 10a, pads 260 formed at a common routing level at the bottom portions of vias are formed first. Dielectric fill 304 is filled to the height where the first routing level is to be formed and is planarized. Via holes are etched and filled with conductive material, for example polysilicon 305. Vias that will reach to the first routing level have now been formed, along with bottom portions of vias that will reach to the second routing level. As shown in Fig. 10b, a conductive material or materials, preferably titanium nitride 306 and tungsten 308, are deposited and first routing level pads 240 are patterned and etched. The titanium nitride etch should be carefully controlled to stop when it reaches polysilicon 305; some overetch can be tolerated.

[0053] Next, as shown in Fig. 10c, a thin etch stop layer 22, for example silicon nitride, is formed over and between pads 240.

[0054] As an alternative, if the thermal budget permits its use, aluminum can be substituted for tungsten in pads 240. In this case, titanium nitride and tungsten can be used in place of polysilicon 305. This may be particularly advantageous, for example, in logic or memory arrays which are formed in the substrate.

[0055] Turning to Fig. 10d, dielectric fill 304 is deposited between and over pads 240, filling to the height at which the second routing level will be formed. Via holes are

etched using a two-step etch process. First a silicon dioxide etch forms the via holes and stops on silicon nitride etch stop layer 22. Next silicon nitride etch stop layer 22 is etched, stopping on the silicon of vias 20 and, in the case of misalignment, silicon dioxide fill 304.

[0056] Conductive material, preferably a layer of titanium nitride and a layer of tungsten, is deposited, filling top portions of the via holes and forming a layer above the dielectric, which is then patterned and etched to form second routing level pads 250. Gaps between pads 250 are filled with dielectric 304, creating structure shown in Fig. 10d.

[0057] In this example, first vias were connected at top ends to the first routing level (pads 240), and second vias were connected at top ends to the second routing level (pads 250.) The bottom portions of the first via holes and the second via holes were etched in the same etch process. The bottom portions of first via holes and the second via holes were filled at substantially the same time.

alternative, when the staggered routing levels are formed above the vias, it may be desirable to form the vias connected to the first routing level and the vias connected to the second routing level completely independently. In this method, pads 260 at a common routing level are formed first. Next dielectric fill 304 is deposited to the height of the first routing level and planarized. Via holes are patterned and etched only for the vias that will connect to the first routing level, then conductive material is deposited, patterned and etched to form first vias 20 and first routing level pads 240. The structure at this point is shown in Fig. 11a.

[0059] Next dielectric fill 304 is deposited to the height of the second routing level and planarized. Via holes are patterned and etched only for the vias that will connect to the second routing level, then conductive material is deposited, patterned and etched to form second vias 30 and second routing level pads 250. The gaps between second

routing level pads 250 are filled by dielectric 304. The resulting structure is shown in Fig. 11b.

[0060] In this example, first vias 20 were connected at top ends to the first routing level (pads 240), and second vias 30 were connected at top ends to the second routing level (pads 250.) The first via holes and the second via holes were etched in separate etch processes, and the first via holes and second via holes were filled at different times.

[0061] The foregoing detailed description has described only a few of the many forms that this invention can take. For this reason, this detailed description is intended by way of illustration, and not by way of limitation. It is only the following claims, including all equivalents, which are intended to define the scope of this invention.